Plotting the Effects of Industrialization: An Interdisciplinary Project

Kasey Vens, *Sylvania Northview High School*

Abstract: *This interdisciplinary project combined a unit on basic data analysis with the social studies curriculum. Students created parallel box plots to compare and contrast a group of industrialized countries with a group of industrializing countries. Students analyzed twelve variables such as literacy rate, infant mortality, and GDP per capita in an effort to better understand the effects of industrialization. Due to the relevance of the data, students were able to have meaningful discussions of the differences in medians and interquartile ranges of the two groups.*

Keywords. *Data analysis, real-world, interdisciplinary, project-based learning*

1 Introduction

The Common Core Standards (CCSSI, 2011) state that students should "understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape." The same initiative outlines eight Standards of Mathematical Practice that emphasize making sense of problems in context and constructing viable arguments. Yet, a common theme of research on statistics in the middle and secondary grades is that it is taught procedurally and superficially. Bakker (2004, p. 64) found that "students generally lack the necessary conceptual understanding to analyze the data." This is a result of two related mistakes teachers often make when addressing statistical concepts with their students. First, many of the datasets teachers ask students to analyze lack context. Textbook and worksheet examples are often nothing more than groups of 10 - 25 numbers. Do they represent quiz scores, calories, miles per gallon, or some other quantity? Supplying students a dataset without any context is an invitation for them to carelessly repeat a procedure. Whether the students are determining a measure of center or producing a graphical display of the data, their results will probably lack any meaningful analysis. Second, the teacher's planned end point for a lesson on statistics is merely to show students how to calculate a measure of center or how to produce a graph as opposed to using those results to encourage students to draw more meaningful conclusions.

Instead of operating in the procedural manner described above, students should study and use statistics to answer questions that provide some insight as to why the data might have been collected in the first place. Friel, O'Connor, and Mamer (2006) propose that these questions should involve students in comparing and contrasting exercises between two datasets and drawing and justifying conclusions. The interdisciplinary project detailed in this paper required students to do both as they attempted to answer the question: *Are industrialized countries statistically distinguishable from underdeveloped countries?*

2 A Conversation and an Idea

During a conversation in our team's common planning time, an unexpected opportunity to mesh topics from social studies and mathematics emerged. Kathy, the social studies teacher, mentioned that during her unit on the Industrial Revolution the class always "looked at" data to see how the more industrialized world differed from other, underdeveloped parts. She aimed to help her students understand the increasing gap in the standard of living between the developed and underdeveloped world caused by industrialization. As a secondary goal, she wanted her students to appreciate the many measurements used to assess a country's standard of living. In Algebra, I would help the students compare and contrast data from two carefully selected groups of countries. In World Studies, Kathy would lead discussions on the same measurements. The results of working from two perspectives were clear: students gained a deeper understanding of the measures of standard of living because they delved deeper into the numbers used to assess it, and they gained a deeper understanding of some basic statistics because of the richer context.

3 Parallel Box Plots

I decided to teach the students to compare the two groups of countries using parallel box plots. Parallel box plots allow for quick, visual comparisons of the medians of and variations within two or more datasets. For our project, one dataset consisted of a sample of 20 industrialized nations (see Table 1 for examples of industrialized nations) and the other dataset consisted of a sample of 20 industrializing nations. When compiling the groups of countries, Kathy and I tried to use many nations with which the students would have some level of familiarity, even if it was only name and general geographical location. The countries we agreed upon make up the columns in the spreadsheets the students compiled, as shown in Table 1.

Table 1: *List of industrialized nations and the variables studied*

	Argentina	Australia	Belgium	Brazil	Canada
GDP Per Capita	14700	41000	37800	10800	39400
Unemployment	7.9	5.2	8.3	6.7	8
Population in Poverty	30	13	15.2	26	9.4
Rate of Inflation	22	2.8	2.3	5	1.8
Fertility Rate	2.3	1.8	1.7	2.2	1.6
Health Expenditures	9.5	8.5	11.8	9	10.9
Population Growth Rate	1	1.1	0.1	1.1	0.8
Birth Rate	17.5	12.3	10.1	17.8	10.3
Median Age	30.5	37.7	42.3	29.3	41
Infant Mortality	10.8	4.6	4.3	21.2	4.9
Life Expectancy	77	81.8	79.5	72.5	81.4
Literacy	97.2	99	99	88.6	99

4 Implementing the Project

4.1 Day One: Researching the Data

In Kathy's social studies class, students were already using the on-line *CIA World Factbook* to research data about countries for other activities. At this site, with a few clicks of the mouse, students can find a wealth of information that could serve as indicators of development for the countries in our project. From all the information available on the site, we compiled a list of 10-12 variables that young adolescents could comprehend and discuss. These make up the row headings in the spreadsheets in Tables 1 and 2. To begin the project, I assigned one or two of the 40 total countries to each student. In the computer lab, the students used the *CIA World Factbook* to research the desired information for each country.

4.2 Day Two: Creating Parallel Box Plots

At the beginning of the second day, the students formed pairs. Then, I randomly selected one pair at a time to choose a variable from the list we had researched the day before, not allowing repetition, until they chose all of the variables. Using the statplot function on the Texas Instruments graphing calculator, each pair of students produced a parallel box plot for one row of the spreadsheet. For example, one pair of students created a box plot of the literacy rates for the sample of 20 industrialized nations. Then, the same pair of students also created a box plot of the literacy rates for the sample of 20 industrializing nations. Ultimately, graphing these two plots on the same number line led to provocative discussions as students used reasoning grounded in both statistics and social studies to explain the similarities and differences between the two box plots.

Before the students progressed to the explanation phase of the project, each pair was responsible for producing a colorful, poster-sized version of their parallel box plot that Kathy would hang in her classroom and refer to for the rest of the year. During this process, the "trace" button of the TI-83/84 series was invaluable. This feature allowed students to move the cursor over a dataset's five-number summary and other important values, including any and all outliers. The cursor may also be moved from one dataset to another in order to make instant comparisons. At this stage, utilizing technology enabled students to quickly draw a parallel box plot, and then transfer it onto poster board accurately. I encountered plenty of teachable moments while the students made their final box plots in this way. For instance, as students determined the scale and label for the x-axis, they began to appreciate, discuss, and ask questions about the criteria used to measure variables like infant mortality, population in poverty, and rate of inflation.

4.3 Day Three/Four: The Graph Is Not the Answer

On the third day, students wrote summaries that highlighted the differences they noticed between the two groups of nations within their own variable, and on the fourth day they presented their findings to the class in a two minute summary. In order to write and speak coherently on their particular variable, each pair of students used what they had learned about the countries in social studies to account for the statistical differences they noticed in the box plots. During these class periods, I found myself repeating the mantra, "once you have your graph done, you are ready to start the project." I hoped this would emphasize the importance of explaining why their graphs looked the way they did.

One of my favorite exercises as I rotated around the room was to quickly cover up the labels on the graph and ask the students which box plot represented which group of countries. Students were often tempted to conclude that the graph with the higher median was always the set of industrialized nations. However, this is not the case for variables such as population growth rate, infant mortality rate, unemployment rate, and the rate of inflation. I asked the students to explain

how they knew which boxplot belonged to which dataset without consulting their spreadsheet, which required the students to make an educated guess as to what effect industrialization might have on their variable.

5 Results

The pairs of students analyzing life expectancy, infant mortality, and median age were all able to correctly identify the presence of better hospitals, health care, and basic living conditions as the most probable causes for differences in the medians for their respective variables. When I asked the pair presenting on life expectancy (Fig. 1) to explain why the graph for the developed nations was so "smashed together," one girl responded by saying all the doctors in those countries are about the same.

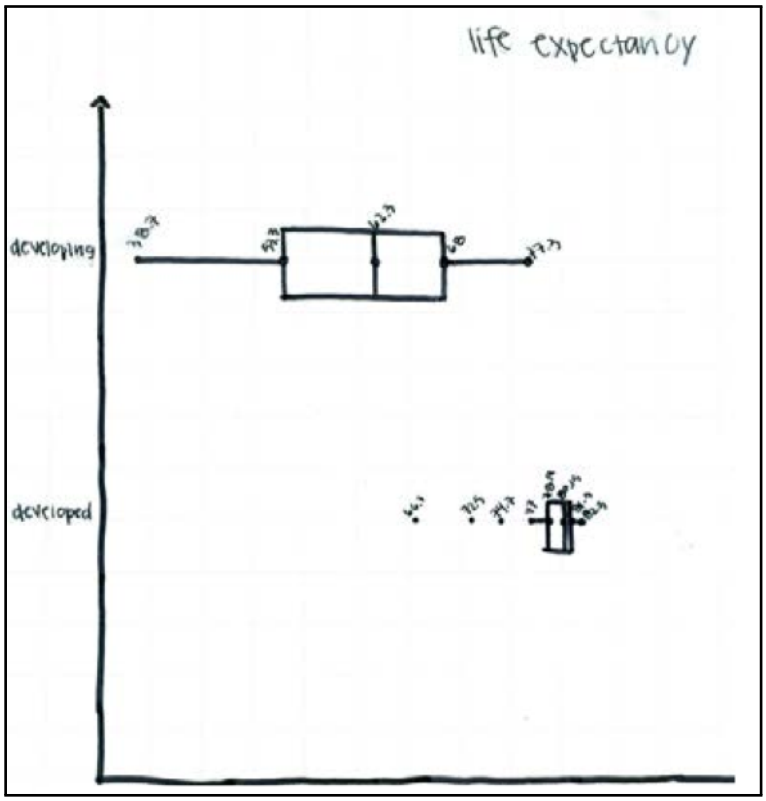

Fig. 1: *Student box plots of life expectancy*

Even though this is clearly an oversimplification, she was beginning to connect a small box (interquartile range) on the graph to the notion of the group's elements showing within-group similarity. Then, a student from the class asked about the outliers for that same group. The students presenting did not have a response. Interestingly, though, another student from the class consulted his spreadsheet and offered the fact that those three outliers were Russia, Brazil, and China, three of the largest countries on the list. He concluded his observation with the suggestion that taking care of everyone in those large countries might be hard. Later, I realized this student was in the GDP per capita group and had already formulated an analogous argument to explain the low annual incomes per capita in China and Brazil. This is an example of a connection that was only possible because a meaningful context was presented and the students were encouraged to go beyond the procedure of simply producing a box plot.

The pairs presenting on percent of population in poverty and unemployment rates (Fig. 2) both

surmised that the medians in the developed nations were lower because industrialization would lead to more job opportunities for the people in those countries.

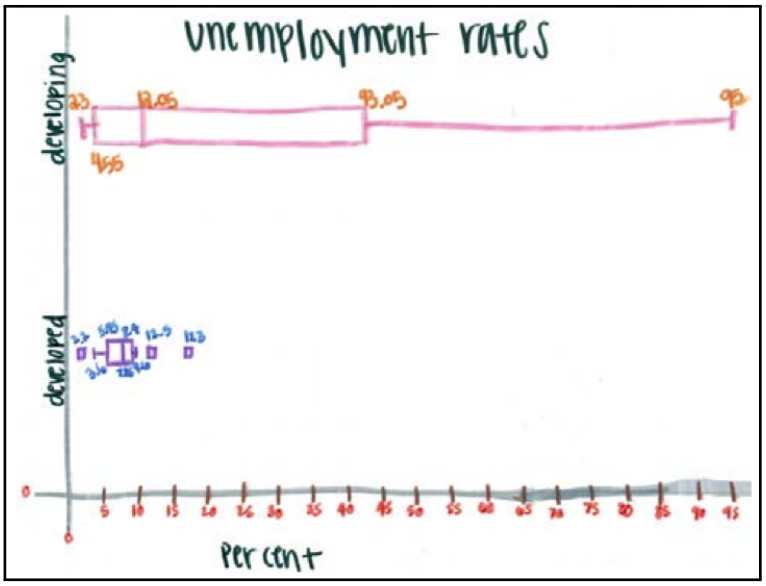

Fig. 2: *Student box plots of unemployment rates*

One student in the unemployment rate group wrote that the extreme difference in spread between the two plots was due to the industrializing nations being a "mainly agricultural economy without factories, offices, and large construction sites to work at...and farming is more unpredictable because of weather." According to Kathy, this student invoked an important concept from social studies. He recognized that a nation's GDP is composed of various sectors. Most often, when a country has not industrialized, the largest portion of its GDP results from agriculture. From my point of view, relating unpredictability to a large spread was equally commendable. The presentation on literacy rates (Fig. 3) was extremely lively.

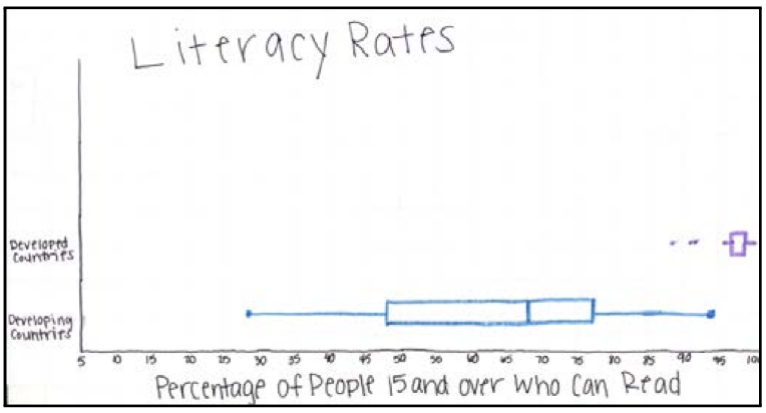

Fig. 3: *Student box plots of literacy rates*

I asked, "Is it fair to say that people in developing countries are stupid?" The presenters rejected this claim by showing countries on the developing list with a larger percentage of the population reading than some countries on the developed list. My intentionally insensitive question was met with a flurry of responses. One of the girls presenting on this topic was born in Nigeria, and she gave three possible explanations. First, she informed us that in some countries education is not free,

girls are not permitted to seek an education, and families are so poor that the children must work to earn money instead of attending school. The room was silent for a second after she finished. I wondered if any of the students felt a greater appreciation for the fact that they are *required* to come to school every day. A humorous exchange occurred when the fertility rate (Fig. 4) pair started to write their summary paragraph.

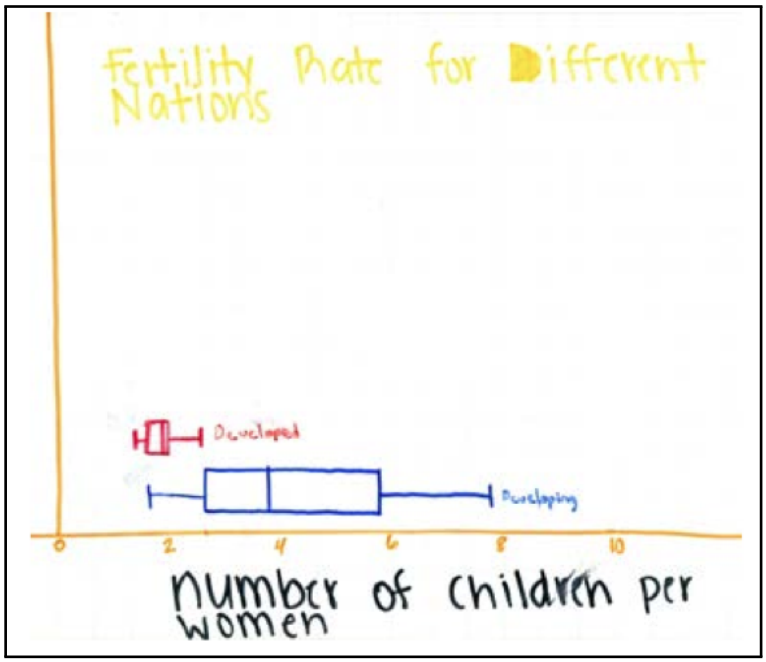

Fig. 4: *Student box plots of fertilty rates*

For two days, while the students researched countries on the Internet and used graphing calculators to find quartiles and draw box plots, the data had been nothing more than numbers on a page. Then, as if struck by lightning, one girl in the pair realized that women in Niger have an average of 7.6 babies in their lifetime. She yelled that she would never have that many children. She and her partner correctly noted that 75% of the developing countries report birth rates higher than any developed country because the lower quartile of the developing sample of nations lay at the max of the developed sample of nations. "Why?" I asked. She hypothesized that in many of those underdeveloped countries a significant number of women did not have careers and "basically stayed home to have babies."

6 Conclusion

In their presentations to the class, most pairs of students were able to explain coherently what their variable measured. More importantly, most pairs of students were able to identify a reasonable explanation for why the two box plots in their graph looked different. The focus of the students' presentations and paragraph summaries rarely deviated from a simple discussion of the differences in medians and interquartile ranges of the two samples of nations. But, we should note that even this is an improvement in using statistics when compared with merely determining these quantities in a meaningless set of data in which no context exists to draw or justify any conclusions.

Over the course of the four days, I enjoyed many opportunities to work one on one with the various pairs of students. With almost every one, I joined a discussion that led to a genuinely teachable moment. In the case of population growth rate, students need to make sense of some negative

data. When students proposed that a negative growth rate indicates the country's population is shrinking, the discussion sounded like one that might be heard in a calculus classroom.

On a final note, this project accomplished something that is nearly impossible to teach directly: some students awakened as conscientious citizens. Whether it was the 15 year-old age cut-off in the literacy rate description or the annual incomes in some countries that were less than their family's cell phone bill each month, the students were seeing the data as something more than simply numbers on a page. Their work shed some light on the drastic difference between the "haves" and the "have nots" on this planet, and in some cases made them appreciate how fortunate everyone in the class is. That is enough to make any social studies teacher proud.

References

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal 3*, 64-83.

Friel, S., O'Connor, W., & Mamer, J. (2006). More than 'Mean median mode' and a bar graph: Whats needed to have a statistical conversation?" In G. Burrill (Ed.), *Thinking and Reasoning with Data and Chance*. Reston, VA: NCTM, 117-137.

Common Core State Standards Initiative (CCSSI) (2010). *Common core standards for mathematics*. (Also available at `http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf`)

Kasey Vens, `kvens@sylvaniaschools.org`, Kasey Vens is pursuing a doctorate in curriculum and instruction from the University of Toledo. He currently serves as assistant principal for curriculum at Sylvania Northview High School. Kasey has thirteen years of teaching experience.